MAT 2379 A
Final Examination

December 12, 2010 **Professor Raluca Balan**
Time: 3 hours

Student Number:_____ **Seat Number:** _____

Family Name: _____ **First Name:** _____

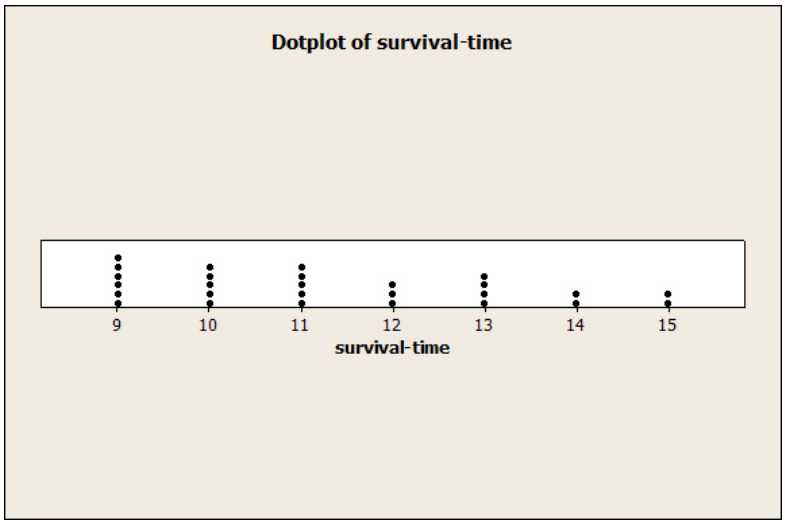
- **This is an open book examination. Only TI 30 calculators are permitted.**
- **Record your answer to each question in the table below. Each question is worth 1 mark.**
- **At the end of the examination, hand in only this page.**

Question	Answer	Question	Answer
1		13	
2		14	
3		15	
4		16	
5		17	
6		18	
7		19	
8		20	
9		21	
10		22	
11		23	
12		24	

****

Professor's use only: Grade=_____/24

1. The following graph is the dotplot of the data which gives the survival time (in weeks) after chemotherapy of 27 terminally-ill cancer patients.



Find the mean ($\bar{x}$) and the median ($\tilde{x}$).

A) $\bar{x} = 11.296, \quad \tilde{x} = 11$ B) $\bar{x} = 12.563, \quad \tilde{x} = 12$
C) $\bar{x} = 10.541, \quad \tilde{x} = 11$ D) $\bar{x} = 11.575, \quad \tilde{x} = 13$
E) $\bar{x} = 10.332, \quad \tilde{x} = 12$

2. A chemical spill occurring in a river affected the fish downstream. It is estimated that 30% of the fish are contaminated, i.e. contain more than 1.5 mg of the chemical per kilogram of body weight. A fisherman has caught 5 fish. What is the probability that at least 2 fish are contaminated?

A) 0.5282 B) 0.4718 C) 0.1631 D) 0.3601 E) 0.8369

3. Let X be the number of smokers in a family composed of a husband and wife. X is a random variable which takes the values 0, 1, 2, with respective probabilities p_0, p_1, p_2. The average number of smokers per family is 0.63. 40% of families contain at least one smoker. Find the probabilities p_0, p_1 and p_2.

A) $p_0 = 0.6, \ p_1 = 0.35, \ p_2 = 0.05$ B) $p_0 = 0.6, \ p_1 = 0.2, \ p_3 = 0.2$
C) $p_0 = 0.6, \ p_1 = 0.17, \ p_2 = 0.23$ D) $p_0 = 0.4, \ p_1 = 0.4, \ p_3 = 0.2$
E) $p_0 = 0, \ \ p_1 = 0.31, \ p_2 = 0.16$

4. The data summarized by the Minitab output below represents the time (in minutes) to the onset of an allergic reaction to a bee sting for 42 patients.

```
Descriptive Statistics: reaction-time

Variable        N  N*    Mean  SE Mean  StDev  Minimum      Q1
reaction-time  42   0  15.248    0.659  4.272    4.100  12.500

Variable       Median       Q3   Maximum
reaction-time  15.050   18.000    27.000
```

Give the interquartile range (IQR) and identify some outliers, if they exist.
A) $IQR = 22.9$, there are no outliers
B) $IQR = 13.9$, there are no outliers
C) $IQR = 14.5$, the value 4.10 is an outlier
D) $IQR = 6.5$, the values 12.50 and 18.00 are outliers
E) $IQR = 5.5$, the values 4.10 and 27.00 are outliers

5. The meteorological data collected for a Canadian city show that the amount of snow during the month of December has a normal distribution with standard deviation $\sigma = 25$ cm. During the last 10 years, this city received an average amount of 82.6 cm of snow, with a standard deviation of 24.1 cm. Find a 96% confidence interval for the average amount of snow received by this city in December.

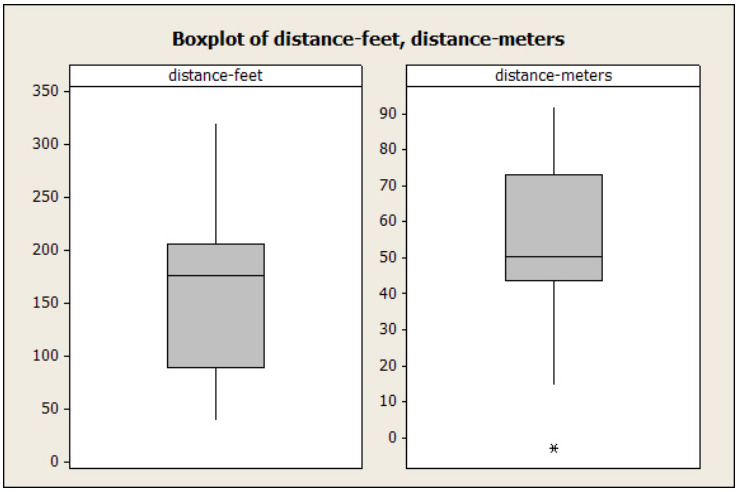
A) [63.6, 101,6] B) [66.4 , 98.8] C) [64.3, 100.9] D) [65.4, 99.8]
E) [67.1, 98.1]

6. Let μ be the average systolic blood pressure of a large population of hypertensive patients who have been taking a certain medication for one month. Using the data supplied by a sample of 12 patients, it was found that a 95% confidence interval for μ is [124.5, 135.1]. Find s, the sample standard deviation.

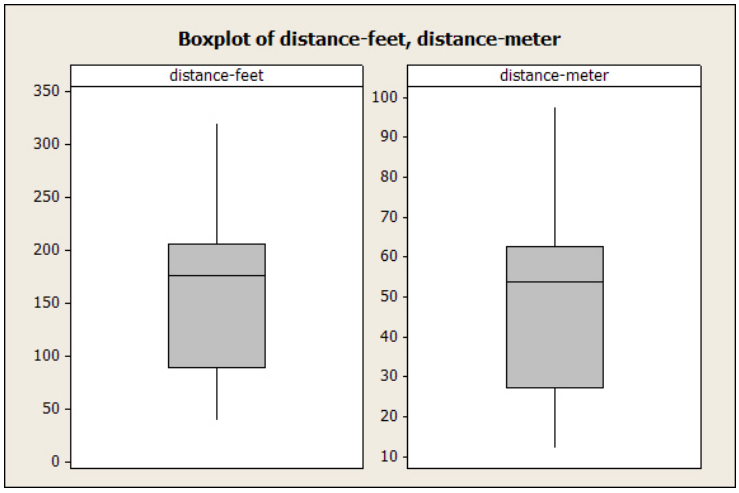
A) 5.30 B) 2.41 C) 18.36 D) 9.37 E) 8.34

7. To study the behavior of morning doves, researchers captured 25 birds, tagged them, released them, and then recorded the flight distance from the point of release to the first stop. The flight distances in feet and meters are recorded in Minitab in columns C1, respectively C2. Which one of the following four pictures gives the correct boxplots of the data in columns C1 and C2? (1 foot=0.3048 meters)

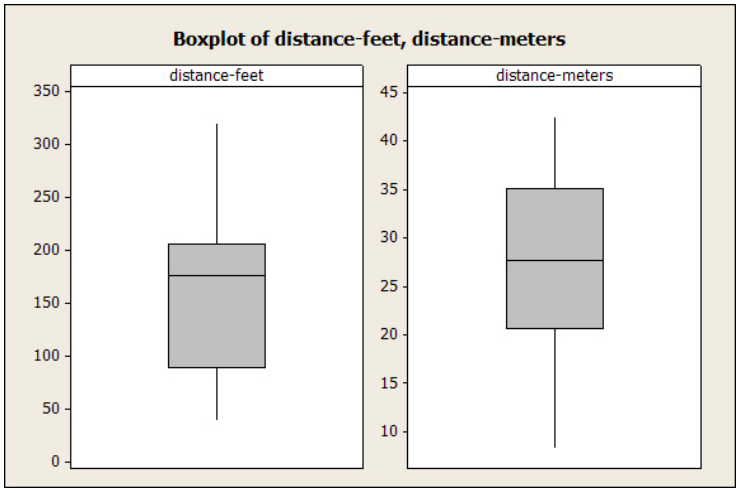
A)



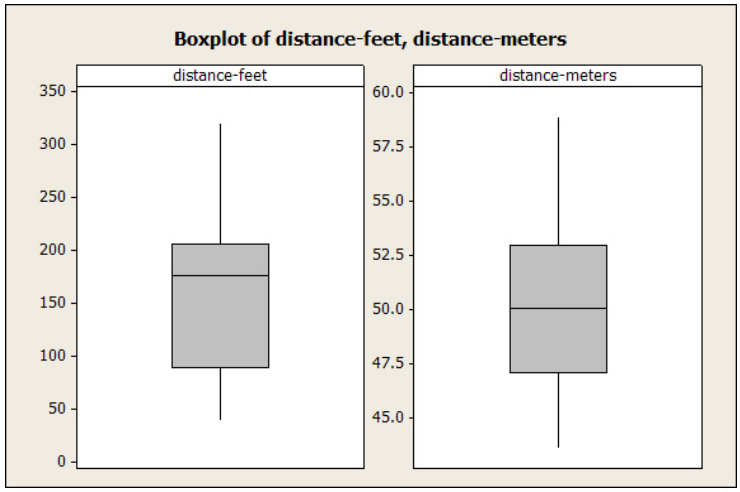
B)



C)



D)



8. A large percentage of the trees in British Columbia have been affected by forest fires or western spruce budworms in the recent years. It is estimated that 75% of the forest remains in healthy condition, 12% has been devastated by a fire but not by budworms, and 5% has been damaged by budworms but not by fires. What is the probability that a randomly chosen tree has been affected by both fires and budworms?

A) 0.60 B) 0.24 C) 0.08 D) 0.07 E) 0.13

9. According to Ontario legislation, passengers aged 13 or older can travel in the front seat of a motor vehicle. The following table gives the extent of injuries and the passenger position for 1000 accidents.

Extent of injury	Front Seat	Back Seat
None	188	70
Minor	232	295
Major	102	75
Death	23	15
Total	545	455

Based on this data, what is the probability of a passenger dying in a motor vehicle accident, given that the passenger was traveling in the front seat? Is death independent of the passenger position?

A) 0.04; yes B) 0.15; yes C) 0.05; no D) 0.042; no E) 0.5; yes.

10. Based on the data given in problem 9, what is the probability that a passenger traveled in the back seat, given that the passenger did not have any injuries?

A) 0.15 B) 0.36 C) 0.27 D) 0.56 E) 0.21

11. The weight of a watermelon is normally distributed with mean 1.5 kg and standard deviation 0.7 kg. Due to overproduction and lack of transportation, a farmer has to discard a quarter of his watermelons. He decides to discard all watermelons with a weight smaller than x_0. What is x_0?

A) 0.93 B) 0.70 C) 1.25 D) 0.50 E) 1.03

12. The mutation of the BRCA gene seems to be related to breast cancer. In a study which involved 393 women, the mutation was found in 14 women. Give a 99% confidence interval for the proportion p of women with a mutation of the BRCA gene.

A) [0.017, 0.053]
B) [0.012, 0.060]
C) [0.014 , 0.057]
D) [0.014, 0.051]
E) [0.02,0.051]

13. Maple syrup is a non-negligible nutritional source of manganese. It is made from the sap of two species of maple: the sugar maple and the black maple. To see if there is a significant difference between the manganese content of these two species of maple, we measured the manganese content (in mg per 100g syrup) for a sample of sugar maples and a sample of black maples. The data is summarized by the Minitab output below:

```
Two-sample T for Sugar Maple - Black Maple

group          N    Mean   StDev  SE Mean
Sugar Maple   20    3.29    0.22    0.049
Black Maple   12    3.55    0.29    0.084
```

Find: (i) a point estimate for $\mu_1 - \mu_2$, where μ_1 and μ_2 are the average manganese contents of the sugar maple, respectively the black maple; (ii) the standard error of $\bar{X}_1 - \bar{X}_2$.

A) (i) -0.07; (ii) 0.101
B) (i) -0.26; (ii) 0.188
C) (i) -0.26; (ii) 0.097
D) (i) -0.26; (ii) 0.101
E) (i) -0.07; (ii) 0.035

14. The Easter tiger salamander is considered extinct in Ontario, but it persists in the prairies and British Columbia. Adults are usually dark with spots of color grey, green or black. In a random sample of 210 tiger salamanders, 60 have grey spots, 80 have green spots, and 70 have black spots. We would like to test the null hypothesis that the three colors of salamander spots are equally likely: $P\{\text{grey}\} = P\{\text{green}\} = P\{\text{black}\} = 1/3$. Give the range of the P-value and the conclusion of the test at level $\alpha = 0.05$.

A) P-value< 0.01; the 3 colors are not equally likely
B) $0.01 <$P-value< 0.02; the 3 colors are not equally likely
C) $0.02 <$P-value< 0.05; the 3 colors are not equally likely
D) $0.05 <$P-value< 0.10; the 3 colors are equally likely
E) P-value> 0.10; the 3 colors are equally likely.

15. A botanist is testing a new tomato fertilizer. He was growing two different groups of 8 plants each, using the standard fertilizer for the first group, and the new fertilizer for the second group. After 70 days, he measured the tomato yield (in kg) for each plant. The data is given in the table below:

Plant	Standard Fertilizer	Plant	New Fertilizer
1	4.76	1	5.60
2	4.25	2	4.98
3	3.98	3	5.12
4	3.44	4	3.86
5	3.87	5	4.56
6	4.78	6	5.76
7	3.99	7	4.21
8	3.21	8	4.05
Mean	4.035	Mean	4.767
Standard deviation	0.56	Standard deviation	0.71

Find a 95% confidence interval for $\mu_1 - \mu_2$, where μ_1 is the average tomato yield per plant using the standard fertilizer, and μ_2 is the average tomato yield per plant using the new fertilizer.

A) [-1.42, -0.04]
B) [-0.96,-0.51]
C) [-1.87, 0.21]
D) [-1.35, -0.31]
E) [-2.44, 0.78]

16. The weight of an adult landlocked Atlantic salmon has a normal distribution with mean 1.4 kg and standard deviation 0.4 kg. A fisherman has caught 4 adult salmons in one day. Calculate:
(i) the probability that a randomly chosen adult salmon has a weight between 1.5 kg and 2.0 kg;
(ii) the probability that the average weight of the 4 salmons is between 1.5 kg and 2.0 kg.

A) (i) 0.3345 (ii) 0.0836
B) (i) 0.0877 (ii) 0.0777
C) (i) 0.3345 (ii) 0.3072
D) (i) 0.3345 (ii) 0.3345
E) (i) 0.3345 (ii) 0.1587

17. The Canadian cow which produces about 18 liters of milk per day has almost disappeared today, being replaced by breeds like Holstein, which produce about 30 liters of milk per day, with 3.25% fat. However, the milk of the Canadian cow is richer in fat and proteins, which is ideal for cheese making. Suppose that a farmer has 6 Canadian cows whose milk has an average level of fat of 4.5%, with a standard deviation of 0.418%. Find a 98% confidence interval for the average percentage of fat found in the milk of Canadian cows.

A) [4.06, 4.94] B) [4.27, 4.73] C) [3.42, 5.57] D) [4.03, 4.97]
E) [3.93, 5.07]

18. A feeding test is conducted on a sample of 24 milking cows to compare two diets: the first diet consists of field wilted alfalfa, the other diet consists of de-watered alfalfa. A sample of 12 cows, randomly selected from the herd, were given the field-wilted alfalfa; another sample of 12 cows were given the de-watered alfalfa. The milk production (in pounds) for each cow is given below:

Field-wilted alfalfa	44	44	56	46	47	38	58	53	49	35	46	30
De-watered alfalfa	35	47	55	29	40	39	32	41	42	57	51	39

Below is the Minitab output which gives the summary of the data:

```
group          N   Mean   StDev   SE Mean
Field-wilted  12  45.50    8.25      2.38
De-watered    12  42.25    8.74      2.52
```

Let μ_1 be the average milk production of cows which are on the field-wilted alfalfa diet, and μ_2 be the average milk production of cows which are on the de-watered alfalfa diet.
(i) Find a 95% confidence interval for $\mu_1 - \mu_2$;
(ii) Test $H_0 : \mu_1 = \mu_2$ against $H_A : \mu_1 \neq \mu_2$ at level $\alpha = 0.05$.

A) (i) $[-6.75, 10.75]$; (ii) 0.10 <P-value<0.20, we do not reject H_0
B) (i) $[-3.97, 10.47]$; (ii) 0.10 <P-value<0.20, we do not reject H_0
C) (i) $[-3.97, 10.47]$; (ii) P-value> 0.20, we do not reject H_0
D) (i) $[-4.79, 11.29]$; ii) P-value<0.05, we reject H_0
E) (i) $[-3.95,\ 9.87]$; (ii) 0.05 <P-value< 0.10, we do not reject H_0.

19. International medical graduates (IMGs) who want to obtain a licence to practice medicine in Ontario have to go through the steps of a highly competitive selection process, despite the fact that they have earned a medical degree in their country of origin. Given the current lack of physicians in Ontario, some people argue that a larger number of IMGs should be given the right to practice. The following table summarizes the attitude towards IMGs of two groups: the first group consist of 50 physicians, the other group consists of 50 non-physicians.

	Physicians	Non-physicians
Larger number of IMGs: yes	30	35
Larger number of IMGs: no	20	15
Total	50	50

Is there enough evidence that the proportions of people who are in favor of giving the right to practice to a larger number of IMGs are different among the physicians, compared to the general population? Justify your answer using a test of hypothesis at level $\alpha = 0.05$. Report the range of the P-value.

A) The proportions are different; $0.01 <$ P-value < 0.02
B) The proportions are different; $0.02 <$ P-value < 0.05
C) The proportions are the same; $0.05 <$ P-value < 0.10
D) The proportions are the same; $0.10 <$ P-value < 0.20
E) The proportions are the same; P-value > 0.20.

20. 16 subjects participated in an experiment to study the effectiveness of a diet program. For each subject, let x be the weight at the beginning of the program, y be the weight at the end of the program, and $d = x - y$ be the difference between the weight before the program and the weight after the program. Suppose that the sample mean and sample variance of the difference data set are $\overline{d} = 0.75$, respectively $s_d^2 = 1.44$. Is there enough evidence that the program was effective in reducing the weight? Justify your conclusion using a test of hypothesis at level $\alpha = 0.05$. Report the range of the P-value.

A) the program was effective, $0.01 <$ P-value < 0.02
B) the program was effective, $0.02 <$ P-value < 0.05
C) the program was not effective, $0.05 <$ P-value < 0.10
D) the program was not effective, $0.10 <$ P-value < 0.20
E) the program was not effective, P-value > 0.20.

21. Crystalline forms of certain chemical compounds are used in various electronic devices. It is often more desirable to have large crystals rather than small ones. In a laboratory study, 14 crystals of the same initial size were allowed to grow for certain periods of time. The following data gives the weight y of the crystal (in grams) and the period x of time (in hours) which was used for each crystal.

Time	Weight	Time	Weight
2	0.08	16	8.4
4	1.12	18	8.81
6	4.43	20	10.81
8	4.98	22	11.16
10	4.92	24	10.12
12	7.18	26	13.12
14	5.57	28	15.04

For this data, we have:

$$\bar{x} = 70, \quad \bar{y} = 7.55, \quad \sum_{i=1}^{14}(x_i - \bar{x})^2 = 910, \quad \sum_{i=1}^{14}(x_i - \bar{x})(y_i - \bar{y}) = 458.12$$

$$\sum_{i=1}^{14}(y_i - \bar{y})^2 = 244.16$$

The time and weight are stored in columns C1, respectively C2. Below is the result of the linear regression analysis produced by Minitab:

```
Regression Analysis: C2 versus C1

The regression equation is C2 = 0.001 + 0.503 C1

Predictor    Coef   SE Coef       T        P
Constant   0.0014    0.5994    0.00    0.998
C1        0.50343   0.03520   14.30    0.000

S = 1.06177   R-Sq = 94.5%   R-Sq(adj) = 94.0%
```

```
Analysis of Variance

Source          DF      SS      MS       F      P
Regression       1  230.63  230.63  204.58  0.000
Residual Error  12   13.53    1.13
Total           13  244.16
```

Assume the linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$. Find a 95% confidence interval for β_1.

A) [-0.426, 0.426]
B) [0.056, 0.950]
C) [0.426, 0.580]
D) [0.405, 0.695]
E) [-0.5,0.5].

22. Continuing with the situation in question 21, which one of the following statements is not true? (There is only one statement which is not true.)

A) Based on the regression model, we expect to have heavier crystals with extra time.
B) The P-value of the test $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$ is smaller than 0.05.
C) The average weight of a crystal which was allowed to grow for 15 hours is 7.546g.
D) The correlation coefficient is -0.9721.

23. Continuing with the situation in question 21, what is the average change in weight of a crystal for every additional 2 hours?

A) 0.002 B) 1.006 C) 0.503 D) 1.007 E) 2.001

24. In Minitab, 50 samples of size $n = 100$ are generated from a normal distribution with the mean $\mu = 5$ and the standard deviation $\sigma = 5$. The confidence intervals for the mean were constructed with the same confidence levels assuming that σ is unknown. The results are given below:

One-Sample T: C1, C2, C3, C4, C5, C6, C7, C8, ...,C50

Variable	N	Mean	StDev	SE Mean	C.I.
C1	100	4.977	5.459	0.546	(3.893, 6.060)
C2	100	5.234	4.826	0.483	(4.276, 6.191)
C3	100	4.596	5.253	0.525	(3.553, 5.638)
C4	100	4.954	5.084	0.508	(3.946, 5.963)
C5	100	4.552	5.356	0.536	(3.489, 5.615)
C6	100	5.394	4.454	0.445	(4.510, 6.278)
C7	100	5.548	5.123	0.512	(4.531, 6.565)
C8	100	5.590	5.418	0.542	(4.515, 6.665)
C9	100	5.414	5.314	0.531	(4.360, 6.469)
C10	100	4.687	6.153	0.615	(3.466, 5.908)
C11	100	6.192	4.826	0.483	(5.234, 7.150)
C12	100	3.938	5.039	0.504	(2.939, 4.938)
C13	100	4.561	4.921	0.492	(3.584, 5.537)
C14	100	4.898	4.976	0.498	(3.911, 5.886)
C15	100	5.858	4.333	0.433	(4.998, 6.718)
C16	100	5.124	4.477	0.448	(4.236, 6.012)
C17	100	5.569	5.364	0.536	(4.505, 6.634)
C18	100	5.507	5.162	0.516	(4.483, 6.531)
C19	100	5.118	5.036	0.504	(4.119, 6.118)
C20	100	4.301	5.163	0.516	(3.277, 5.326)
C21	100	4.799	4.699	0.470	(3.867, 5.732)
C22	100	5.162	4.515	0.452	(4.266, 6.058)
C23	100	4.868	4.577	0.458	(3.960, 5.776)
C24	100	5.701	5.086	0.509	(4.692, 6.710)
C25	100	5.386	5.079	0.508	(4.378, 6.394)
C26	100	4.938	4.725	0.472	(4.000, 5.875)
C27	100	5.215	4.575	0.457	(4.307, 6.122)
C28	100	4.517	4.087	0.409	(3.706, 5.328)
C29	100	5.555	5.169	0.517	(4.530, 6.581)
C30	100	5.182	5.634	0.563	(4.064, 6.300)
C31	100	4.596	4.366	0.437	(3.729, 5.462)
C32	100	5.290	4.635	0.463	(4.371, 6.210)
C33	100	4.534	4.716	0.472	(3.598, 5.469)
C34	100	4.815	5.397	0.540	(3.745, 5.886)
C35	100	4.934	4.873	0.487	(3.967, 5.901)

C36	100	4.599	5.005	0.500	(3.606, 5.592)
C37	100	4.306	4.657	0.466	(3.382, 5.230)
C38	100	4.581	5.056	0.506	(3.578, 5.585)
C39	100	5.190	4.597	0.460	(4.278, 6.102)
C40	100	4.797	4.969	0.497	(3.811, 5.783)
C41	100	6.329	5.054	0.505	(5.326, 7.332)
C42	100	4.883	4.820	0.482	(3.927, 5.840)
C43	100	5.635	4.951	0.495	(4.652, 6.617)
C44	100	5.534	4.863	0.486	(4.569, 6.498)
C45	100	5.109	5.096	0.510	(4.098, 6.121)
C46	100	4.628	4.611	0.461	(3.713, 5.543)
C47	100	4.993	4.869	0.487	(4.027, 5.959)
C48	100	4.406	4.712	0.471	(3.471, 5.341)
C49	100	4.467	4.873	0.487	(3.500, 5.434)
C50	100	5.422	5.133	0.513	(4.403, 6.440)

Estimate the confidence level of the intervals based on the results produced by Minitab.

A) 94% B) 99% C) 80% D) 90%
E) 70%.